

(19)日本国特許庁 (J P)

(12)公開特許公報 (A)

(11)特許出願公開番号

特開平10-55194

(43)公開日 平成10年(1998)2月24日

(51) Int.CI.	識別記号	府内整理番号	F I	技術表示箇所	
G10L 3/00	531	G10L 3/00	531	N	
	561			R	
	571			G	
				571	H

審査請求 未請求 請求項の数12 O L (全10頁)

(21)出願番号 特願平8-210196
(22)出願日 平成8年(1996)8月8日

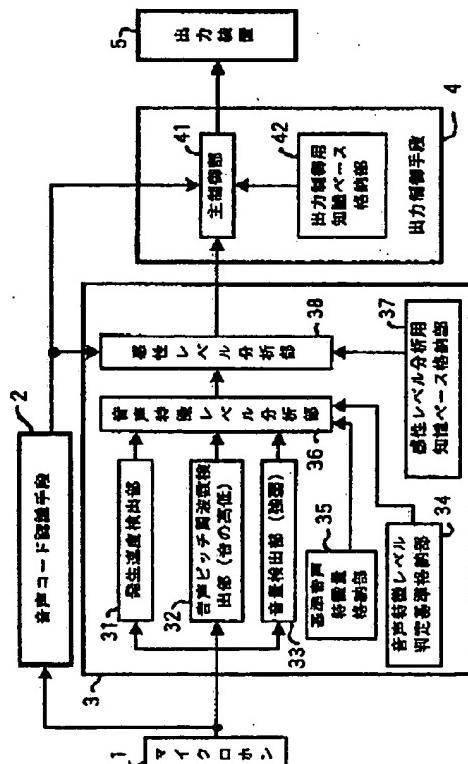
(71)出願人 000001889
三洋電機株式会社
大阪府守口市京阪本通2丁目5番5号
(72)発明者 武田 昭二
大阪府守口市京阪本通2丁目5番5号 三
洋電機株式会社内
(72)発明者 山本 英人
大阪府守口市京阪本通2丁目5番5号 三
洋電機株式会社内
(74)代理人 弁理士 安富 耕二 (外1名)

(54)【発明の名称】音声制御装置と音声制御方法

(57)【要約】

【課題】 音声制御を行う場合に、話者の意図にあった細かな制御ができない。話者は、人と会話する感覚で機械に命令できない。

【解決手段】 話者からの音声情報に付随する状態情報を抽出する状態情報抽出手段3を具備する。そして、人間の状態(感情、体調等)による音声指示の変化を適格に判断する。単なる音声命令だけでなく、その状態を加味して出力装置を制御する。



【特許請求の範囲】

【請求項 1】 音声情報を入力し、その音声情報に対応するコード情報を認識するコード認識手段(2)と、前記音声情報に付随する話者の状態情報を抽出する状態情報抽出手段(3, 3')と、前記コード認識手段(2)の出力と前記状態情報抽出手段(3, 3')の出力に基づいて出力装置(5)を制御する出力制御手段(4)とを備えたことを特徴とする音声制御装置。

【請求項 2】 前記状態とは、話者の体調状態であることを特徴とする請求項 1 記載の音声制御装置。

【請求項 3】 前記状態とは、話者の感情状態であることを特徴とする請求項 1 記載の音声制御装置。

【請求項 4】 音声情報を入力し、その音声情報に対応するコード情報を認識するコード認識手段(2)と、前記音声情報に付随する話者の状態情報を抽出する状態情報抽出手段(3, 3')と、

前記コード認識手段(2)の出力と前記状態情報抽出手段(3, 3')の出力に基づいて、出力装置(5)で出力されるメッセージ出力の内容を変更する出力制御手段(4)とを備えたことを特徴とする音声制御装置。

【請求項 5】 前記出力装置は、音声メッセージ出力機能を備えたことを特徴とする請求項 4 記載の音声制御装置。

【請求項 6】 前記出力装置は、映像メッセージ出力機能を備えたことを特徴とする請求項 4 記載の音声制御装置。

【請求項 7】 入力した音声情報に対応するコード情報を認識し、このコード情報に対応して出力装置(5)を制御する音声制御装置の音声制御方法において、前記音声情報から話者の状態を検出し、この検出した話者の状態と、前記コード情報に基づいて前記出力装置(5)を制御することを特徴とする音声制御方法。

【請求項 8】 入力した音声情報に対応するコード情報を認識すると共に、前記音声情報から話者の状態を検出し、この検出した話者の状態と前記コード情報とに対応して出力装置(5)を制御する音声制御装置の音声制御方法において、

話者の状態を検出出来なかった時は、直前に検出した状態として制御することを特徴とする音声制御方法。

【請求項 9】 入力した音声情報に対応するコード情報を認識すると共に、前記音声情報から話者の状態を検出し、この検出した話者の状態と前記コード情報とに対応して出力装置(5)を制御する音声制御装置の音声制御方法において、

話者の状態を検出出来なかった時は、平常状態として制御することを特徴とする音声制御方法。

【請求項 10】 入力した音声情報に対応するコード情報を認識すると共に、前記音声情報から話者の状態を検出し、この検出した話者の状態と前記コード情報とに対

10

20

30

40

50

応して出力装置(5)を制御する音声制御装置の音声制御方法において、

話者の状態を検出出来なかった時は、非常状態として制御することを特徴とする音声制御方法。

【請求項 11】 入力した音声情報に対応するコード情報を認識すると共に、前記音声情報から話者の状態を検出し、この検出した話者の状態と前記コード情報とに対応して出力装置(5)を制御する音声制御装置の音声制御方法において、

前記出力装置(5)で前記コード情報と話者の状態とに応じて、この話者を気遣うメッセージを出力することを特徴とする音声制御方法。

【請求項 12】 入力した音声情報に対応するコード情報を認識すると共に、前記音声情報から話者の状態を検出し、この検出した話者の状態と前記コード情報とに対応して、出力装置(5)で話者に対してメッセージを出力する音声制御装置の音声制御方法において、このメッセージに対する話者の反応を検出して、この検出結果に応じて出力装置(5)を制御することを特徴とする音声制御方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、声により出力装置を制御する音声制御装置に関する。本発明は、音声を入力して、その音声の指示によって、出力装置を制御する音声制御装置及びその音声制御方法に関する。つまり、制御入力を声で行うマン・マシーンインターフェスとして使用される音声制御方法に関する。

【0002】

【従来の技術】 音声制御装置は、音声を入力し、この声を音声認識し、この声の内容に応じて出力装置を制御している。

【0003】

【発明が解決しようとする課題】 しかし、処理が画一的であり、指令を出す人間側に感情や体調の変化がある場合には、かならずしも人間にとって満足する結果が得られないことがある。ところで、人間の感情や体調等を、発汗量・脈拍数・体温・血圧・音声から検出し、この検出した感情や体調に応じて、機器を制御することは、知られている。本発明は、人間の状態（感情、体調等）も加味した音声制御を提供するものである。

【0004】 又、本発明は、人間の状態（感情、体調等）も加味した音声制御を簡単に実現するものである。又、本発明は、人間の状態（感情、体調等）も加味した音声指示を的確に行うことが出来る音声制御を提供するものである。

【0005】

【課題を解決するための手段】 本発明は、音声制御を行う場合に、人間の状態（感情、体調等）も加味して音声制御を行うことを特徴とする。本発明は、音声情報を入

力し、その音声情報に対応するコード情報を認識するコード認識手段(2)と、前記音声情報に付随する話者の状態情報を抽出する状態情報抽出手段(3,3')と、前記コード認識手段(2)の出力と前記状態情報抽出手段(3,3')の出力に基づいて出力装置(5)を制御する出力制御手段(4)とを備えたことを特徴とする。

【0006】また、本発明は、音声情報を入力し、その音声情報に対応するコード情報を認識するコード認識手段(2)と、前記音声情報に付随する話者の状態情報を抽出する状態情報抽出手段(3,3')と、前記コード認識手段(2)の出力と前記状態情報抽出手段(3,3')の出力に基づいて、出力装置(5)で出力されるメッセージ出力の内容を変更する出力制御手段(4)とを備えたことを特徴とする。

【0007】また、本発明は、入力した音声情報に対応するコード情報を認識し、このコード情報に対応して出力装置(5)を制御する音声制御装置の音声制御方法において、前記音声情報から話者の状態を検出し、この検出した話者の状態と、前記コード情報に基づいて前記出力装置(5)を制御することを特徴とする。また、本発明は、入力した音声情報に対応するコード情報を認識すると共に、前記音声情報から話者の状態を検出し、この検出した話者の状態と前記コード情報とに応じて出力装置(5)を制御することを特徴とする。また、本発明は、入力した音声情報に対応するコード情報を認識すると共に、前記音声情報から話者の状態を検出し、この検出した話者の状態と前記コード情報とに応じて出力装置(5)を制御することを特徴とする。

【0008】また、本発明は、入力した音声情報に対応するコード情報を認識すると共に、前記音声情報から話者の状態を検出し、この検出した話者の状態と前記コード情報とに応じて出力装置(5)を制御することを特徴とする。また、本発明は、入力した音声情報に対応するコード情報を認識すると共に、前記音声情報から話者の状態を検出し、この検出した話者の状態と前記コード情報とに応じて出力装置(5)を制御することを特徴とする。

【0009】また、本発明は、入力した音声情報に対応するコード情報を認識すると共に、前記音声情報から話者の状態を検出し、この検出した話者の状態と前記コード情報とに応じて出力装置(5)を制御することを特徴とする。また、本発明は、入力した音声情報に対応するコード情報を認識すると共に、前記音声情報から話者の状態を検出し、この検出した話者の状態と前記コード情報とに応じて出力装置(5)を制御することを特徴とする。

【0010】また、本発明は、入力した音声情報に対応するコード情報を認識すると共に、前記音声情報から話者の状態を検出し、この検出した話者の状態と前記コード情報とに応じて、この話者を気遣うメッセージを出力することを特徴とする。

【0011】また、本発明は、入力した音声情報に対応するコード情報を認識すると共に、前記音声情報から話者の状態を検出し、この検出した話者の状態と前記コード情報とに応じて、出力装置(5)で話者に対してメッセージを出力することを特徴とする。

て、このメッセージに対する話者の反応を検出して、この検出結果に応じて出力装置(5)を制御することを特徴とする。

【0012】

【発明の実施の形態】以下、本発明の第1実施例について説明する。図1は、本発明の第1実施例の音声制御装置を示すブロック図である。図2は、この第1実施例の音声制御装置の処理動作を説明するための説明図である。

【0013】この図2は、通常の喋り方と怒った喋り方での音声情報の特徴量の違いを示し、図2(a)は音声波形、図2(b)は音量特性、図2(c)は音声ピッチ周波数特性を示している。図1において、マイクロホン1は、音声を電気信号に変換する。音声情報(声)をこのマイクロホン1から入力し、その音声情報に対応する音声コード情報を、音声コード認識手段2が音声認識する。

【0014】この認識された音声コードは、状態情報抽出手段3及び出力制御手段4に出力される。ところで、状態情報抽出手段3は、例えば、音声認識分野で知られているHMM(HiddenMarkovModel)法と称する音声の状態遷移の統計処理を応用して感情の種類とその値(程度)を導き出す。この手法については、書籍「HiddenMarkovModelsforSpeechRecognition」(出版:Edinburg University Press, 著者:X.D.HUNG他, 1990年)で開示されている。

【0015】例えば、通常の喋り方の「ur usai」と怒った喋り方の「ur usai」の音声情報の特徴量の違いを示すと、図2に示すようになる。図2の左側が通常の喋り方での「うるさい」の音声で、その右側が怒った喋り方での「うるさい」の音声である。ここで、右側の波形の方が、音量が大きく、発生時間が短くなっている、発生速度が速く、かつ、音声ピッチ周波数が高いことがわかる。

【0016】この状態情報抽出手段3の発生速度検出部31は、マイクロホン1から入力した音声情報の発生速度を検出する。この状態情報抽出手段3の音声ピッチ(音の高低)周波数検出部32は、マイクロホン1から入力した音声情報の音声ピッチ周波数を検出する。この状態情報抽出手段3の音量検出部33は、マイクロホン1から入力した音声情報の音量を検出する。

【0017】基準音声特徴量格納部35は、音声分析する場合の基準となる音声特徴量(発生速度、音声ピッチ周波数、音量)が記憶されている。この場合、基準音声特徴量として、話者の音声特徴量を基準として格納するケースと、多くの人の音声特徴量の平均(標準値)を基準として格納するケースとが、考えられる。この実施例の基準音声特徴量格納部35では、2つのメモリエリアを持ち、本実施例では、両方に対応できるものとした。

【0018】つまり、メモリエリアの一方に予め使用者(話者)が普通に喋ったときの音声特徴量(発生速度、音

声ピッチ周波数、音量)を基準値として記憶できるものとする。そして、他方のメモリエリアには、予め多くの人の音声特徴量の平均(標準値)を記憶しておく。そして、使用されるメモリエリアは、この音声制御装置と組み合わされる出力装置又は話者(使用者)の好みに応じて、選択出来るようにした。

【0019】ここで、一般的な音声特徴量(発生速度、音声ピッチ周波数)の一般的な値について述べる。まず、発生速度については、単位時間あたりに発生されたモーラ数より検出する。このモーラとは、「拍」と訳され、基本的には、かな文字と同じである。

【0020】このモーラでは、拗音(「きや」「じょ」等)、促音(「っ」)、撥音(「ん」)も1モーラと数える。また、このモーラでは、長音(「そうじ」の「そう」、「テープ」の「テー」)などは2モーラとなる。なお、モーラ数の計測は、音声認識で一般的に使用されている。発生速度を「ゆっくり」、「普通」、「早口」との3段階に大別すると、

ゆっくり : 7モーラ以下/秒

普通 : 8~11モーラ/秒

早口 : 12モーラ以上/秒

が一般的に感じるものである。

【0021】また、音声ピッチは、音声信号を線形予測

発生速度レベル : 早口/普通/ゆっくり

音声ピッチ周波数レベル : 高い/普通/低い

音量特徴レベル : 大きい/普通/小さい

つまり、音声特徴レベル判定基準格納部34には、このレベルに分類するための【発生速度レベル基準】と【音声ピッチ周波数レベル基準】、【音量特徴レベル基準】が格納されている。

【0024】発生速度を「早口/普通/ゆっくり」の3レベルに分類するための【発生速度レベル基準】について説明する。この【発生速度レベル基準】は、入力された音声の発生速度と、基準音声特徴量格納部35に格納された標準の発生速度とを比べ、この比較から「早口/普通/ゆっくり」の3レベルを求めるための基準である。例えば、ここでは、以下のものである。

【0025】

30%以上速ければ : 早口

30%以上遅ければ : ゆっくり

それ以外であれば : 普通

また、音声ピッチを「高い声/普通の声/低い声」の3レベルに分類するための【音声ピッチ周波数レベル基準】が格納される。

【0026】この【音声ピッチ周波数レベル基準】は、入力された音声の音声ピッチと、基準音声特徴量格納部35に格納された標準の音声ピッチとを比べ、この比較から「高い声/普通の声/低い声」の3レベルを求めるための基準である。例えば、ここでは、以下のものである。

分析した後の残差信号の相関から求める。この方法は、書籍「図解、音声デバイス活用の実際」P2~5、(中田和男、三瓶徹共編、オーム社、1982年発行)等で公知である。例えば、日本人の男性成人、女性成人、子供の平均の音声ピッチ周波数は、

日本人の男性成人の平均 : 100~125Hz

日本人の女性成人の平均 : 250~300Hz

日本人の子供の平均 : 125~250Hz

である。

【0022】使用者が男性で、予め基準音声特徴量格納部35に、この使用者の音声特徴量(発生速度、音声ピッチ周波数、音量)を格納する場合、この男性が標準的であれば、音声ピッチ周波数は「100~125Hz」の範囲内となり、発生速度は、「8~11モーラ/秒」の範囲内となる。音声特徴レベル判定基準格納部34には、音声特徴レベル判定基準が格納される。

【0023】この音声特徴レベル判定基準格納部34は、各音声特徴レベル(発生速度レベル、音声ピッチ周波数レベル、音量特徴レベル)を求めるためのものであり、【発生速度レベル】と【音声ピッチ周波数レベル】、【音量特徴レベル】を以下のように3段階とする。

発生速度レベル : 早口/普通/ゆっくり

音声ピッチ周波数レベル : 高い/普通/低い

音量特徴レベル : 大きい/普通/小さい

50%以上高ければ : 高い声

50%以上低ければ : 低い声

それ以外であれば : 普通の高さの声

又、音量を「大きな声/普通の声/小さい声」の3レベルに分類するための【音量特徴レベル基準】が格納される。

【0027】この【音量特徴レベル基準】は、入力された音声の音量と、基準音声特徴量格納部35に格納された標準の音量とを比べ、この比較から「大きな声/普通の声/小さい声」の3レベルを求めるための基準である。例えば、ここでは、以下のものである。

3dB以上高ければ : 大きい声

3dB以上低ければ : 小さい声

それ以外であれば : 普通の大きさの声

尚、音量の比較は、母音部分について比較している。

【0028】このように、音声特徴レベル判定基準格納部34及び基準音声特徴量格納部35に格納されている情報と、各検出部31、32、33からの出力を基に音声特徴レベル分析部36で音声情報が分析される。状態情報抽出手段3では、まず、発生速度検出部31で音声情報の発生速度を検出し、音声ピッチ周波数検出部32で音声ピッチ周波数を検出し、音量検出部33で音量を検出する。

【0029】そして、音声特徴レベル分析部36では、

発生速度検出部 3 1 からの発生速度レベルを分析する。具体的には、まず、検出した発生速度と基準音声特微量格納部 3 5 の標準発生速度との比を求める。次に、この発生速度比と、音声特徴レベル判定基準格納部 3 4 の発生速度レベル判定基準に基づいて、発生速度レベルを分析する。

【0030】また、音声特徴レベル分析部 3 6 では、音声ピッチ周波数検出部 3 2 からの音声ピッチ周波数レベルを分析する。具体的には、まず、検出した音声ピッチ周波数と基準音声特微量格納部 3 5 の基準音声ピッチ周波数との比を求める。次に、音声ピッチ周波数比と音声特徴レベル判定基準格納部 3 4 の音声ピッチ周波数レベル判定基準に基づいて音声ピッチ周波数レベルを分析する。

【0031】また、音声特徴レベル分析部 3 6 では、音量検出部 3 3 からの音量特徴レベルを分析する。具体的には、まず、検出した音量と基準音声特微量格納部 3 5 の標準音量との比を求める。次に、音量比と音声特徴レベル判定基準格納部 3 4 の音量特徴レベル判定基準に基づいて、音量特徴レベルを分析する。

【0032】そして、感性レベル分析用知識ベース格納部 3 7 は、各音声特徴レベルと音声コードとによって、感性レベルを判定するための条件（ルール）を設定しておく。ここで、感性レベルとは、感性の種類と程度を意味する。感性レベル分析部 3 8 では、音声特徴レベル分析部 3 6 からの出力及び音声コード認識手段 2 からの出力を入力し、音声コードと、発生速度レベル、音声ピッチ周波数レベル、音量特徴レベルから、感性レベル分析用知識ベース格納部 3 7 に格納されている感性レベル分析用知識ベースに基づいて、感性レベルを分析する。

【0033】出力制御手段 4 は、感性レベル分析部 3 8 からの出力及び音声コード認識手段 2 からの出力を入力し、各種の出力装置 5 を制御する。出力制御手段 4 は、主制御部 4 1 及び出力制御用知識ベース格納部 4 2 を有している。出力制御用知識ベース格納部 4 2 は、感性レベル分析部 3 8 からの出力及び音声コード認識手段 2 からの出力によって如何なる出力を行えばよいかを判断する基準を格納する条件マップを格納したものである。

【0034】主制御部 4 1 は、感性レベル分析部 3 8 からの出力及び音声コード認識手段 2 からの出力を出力制御用知識ベース格納部 4 2 で照合した後、その条件によって出力装置 5 の制御情報を決定し、その制御情報を出力する。出力装置 5 は、音声制御装置により制御される機器であり、制御しようとする制御対象の機器（産業用機器、民生用機器等）のオン・オフ制御または所定の特性を変化させる制御により、出力が変化する。

【0035】また、出力装置 5 は、メッセージ出力機能を有し、そのメッセージを変更するものであってもよ

い。図 3 は、この第 1 実施例の音声制御装置の状態情報抽出手段 3 が行う制御処理を示すフローチャートである。まず、ステップ S 1 で発生速度検出部 3 1 で音声情報の発生速度を検出する。

【0036】ステップ S 2 で音声ピッチ周波数検出部 3 2 によって音声ピッチ周波数を検出する。ステップ S 3 で音量検出部 3 3 によって音量を検出する。ステップ S 4 で発生速度検出部 3 1 から入力した発生速度レベルを分析する。ここでは、検出した発生速度と基準音声特微量格納部 3 5 の標準発生速度との比を求め、発生速度比と音声特徴レベル判定基準格納部 3 4 の発生速度レベル判定基準に基づいて、発生速度レベルを分析する。

【0037】また、ステップ S 5 で音声ピッチ周波数検出部 3 2 から入力した音声ピッチ周波数レベルを分析し、検出した音声ピッチ周波数と基準音声特微量格納部 3 5 の基準音声ピッチ周波数との比を求め、音声ピッチ周波数比と音声特徴レベル判定基準格納部 3 4 の音声ピッチ周波数レベル判定基準に基づいて音声ピッチ周波数レベルを分析する。

【0038】そして、ステップ S 6 で音量検出部 3 3 から入力した音量特徴レベルを分析し、検出した音量と基準音声特微量格納部 3 5 の標準音量との比を求め、音量比と音声特徴レベル判定基準格納部 3 4 の音量特徴レベル判定基準に基づいて、音量特徴レベルを分析する。更に、ステップ S 7 で感性レベル分析部 3 8 で、音声コードと、発生速度レベル、音声ピッチ周波数レベル、音量特徴レベルから、感性レベル分析用知識ベースに基づいて、感性レベル（感性の種類と程度）を分析する。

【0039】このように、感性レベル分析部 3 8 では、入力された音声コード認識手段 2 からの音声コードと、音声特徴レベル分析部 3 6 で処理された感性レベルとを組合わせ、感性レベル分析用知識ベース格納部 3 7 に格納された感性レベル分析用知識ベース（条件マップ）に基づいて、感性レベルを検出する。感性レベル分析用知識ベース格納部 3 7 に格納された感性分析用知識ベースとしては、予め、普通の喋り方や怒ったときの喋り方等、感性レベルを変えて入力したときの各音声特徴レベルを、この感性レベルごとにデータベースとして記憶しておく。

【0040】そして、音声特徴レベル分析部 3 6 で処理された後、入力された音声の各音声特徴レベルが、データベース内の感性レベルごとの各音声特徴レベルとを照合し、最も一致度の高い感性レベルを検出結果とする。次に、この音声処理装置での処理の事例について説明する。

【事例 1】

【0041】

【表 1】

9

10

入力音声コード	感性レベル分析用知識ベース	制御用知識ベース
「うるさい」	入力音量が大きいとき : 不満度が大きい	不適度が大きい : 機械の出力音を止める
	入力音量が小さいとき : 不満度が小さい	不適度が小さい : 機械の出力音量を小さくする

【0042】この表1の事例は、TVの音量調整に使用する最も単純な例である。ここでは、単純に、1種類の音声特徴レベルから、感性レベルを検出している。つまり、状態情報抽出手段3は、大声の「うるさい」が入力されると、「不満度が大」の感性レベル情報を出力する。

【0043】そして、出力制御手段4は、「不満度が大」の「うるさい」が入力されると、出力装置の消音を

行う。また、状態情報抽出手段3は、小声の「うるさい」が入力されると、「不満度が小」の感性レベル情報を出力する。そして、出力制御手段4は、「不満度が小」の「うるさい」が入力されると、出力装置の出力音量を下げる。

【0044】[事例2]

【0045】

【表2】

入力音声コード	感性レベル分析用知識ベース	制御用知識ベース
「もっと、明るく」	音声ピッチ周波数が高いとき : 要望の程度が大きい	要望の程度が大きい : 非常に明るくする 【2段階明るく】
	音声ピッチ周波数が低いとき : 要望の程度が小さい	要望の程度が小さい : 少し明るくする 【1段階明るく】

【0046】この表2の事例は、TVの画面の明るさ又は照明の明るさ等の調整に使用する最も単純な例である。ここでも、単純に、1種類の音声特徴レベルから、感性レベルを検出している。つまり、状態情報抽出手段3は、音声ピッチ周波数の高い「もっと、明るく」が入力されると、「要望の程度大」の感性レベル情報を出力する。

【0047】そして、出力制御手段4は、「要望の程度大」と「もっと、明るく」が入力されると、出力装置5

の明るさのレベルを2段上げる。また、状態情報抽出手段3は、音声ピッチ周波数の低い「もっと、明るく」が入力されると、「要望の程度小」の感性レベル情報を出力する。そして、出力制御手段4は、「要望の程度小」と「もっと、明るく」が入力されると、出力装置5の明るさのレベルを1段上げる。

【0048】[事例3]

【0049】

【表3】

入力音声コード	感性レベル分析用知識ベース	制御用知識ベース
「止まれ」	音声速度が遅いとき : 危機感が少ない	危機感が少ない : 機械を徐々に減速して停止させる
	音声速度が早いとき : 危機感が強い	危機感が強い : 危機を急停止する

【0050】この表3の事例は、リモコンキーの速度制御を行う単純な例である。ここでも、単純に、1種類の音声特徴レベルから、感性レベルを検出している。つまり、状態情報抽出手段3は、音声速度の遅い「止まれ」が入力されると、「危機感小」の感性レベル情報を出力する。

【0051】そして、出力制御手段4は、「危機感小」と「止まれ」が入力されると、出力装置5を徐々に停止

させる。また、状態情報抽出手段3は、音声速度が速い「止まれ」が入力されると、「危機感大」の感性レベル情報を出力する。そして、出力制御手段4は、「危機感大」と「止まれ」が入力されると、出力装置5を緊急停止させる。

【0052】[事例4]

【0053】

【表4】

11

12

入力音声コード	感性レベル分析用知識ベース	制御用知識ベース
「ちがう」	音声ピッチ周波数が高く、発生速度が速く、音量が大きいとき ：不満度が高い	不満度が高い ：大変緊張した様子の「大変申し訳ありません」の音声出力をを行い、UNDOを行う。
	上記以外の場合 ：不満度が低い	不満度が低い ：平常時の様子の「申し訳ありません」の音声出力をを行い、UND Oを行う。

【0054】この表4の事例は、コンピュータ等の文書処理における「UNDO機能（取り消し機能、元に戻す機能）」に関する。つまり、状態情報抽出手段3は、音声ピッチ周波数が高く、且つ、発生速度が速く、且つ、大声の「ちがう」が入力されると、「不満度が高い」の感性レベル情報を出力する。

【0055】そして、出力制御手段4は、「不満度が高い」と「ちがう」が入力されると、出力装置5において、今まで処理していた動作を止め、元の状態に戻すUNDO機能を実行するとともに、「大変申し訳ありません」の録音メッセージデータを音声出力する。また、状態情報抽出手段3は、上記以外の場合の「ちがう」が入力されると、「不満度が低い」の感性レベル情報を出力する。

【0056】そして、出力制御手段4は、「不満度が低い」と「ちがう」が入力されると、出力装置5において、今まで処理していた動作を止め、UNDO機能を実行するとともに、画面に「申し訳ありません」の表示を行う。このように、本実施例の音声制御装置は、音声情報を入力し、その音声情報に対応する音声コード情報を認識する音声コード認識手段2と、音声情報に付随する感性情報を抽出する状態情報抽出手段3と、音声コード認識手段2の出力と状態情報抽出手段3との出力に基づいて出力装置5の出力を制御する出力制御手段4とを具備している。

【0057】出力制御手段4は、状態情報抽出手段3からの出力と音声コード認識手段2の出力とに応じて如何なる出力をすればよいかを、出力制御知識ベース部42で照合した後、その照合結果である制御情報を出力装置5に出力する。尚、上記実施例では、感性の種類としては、不満度、要望度、危機感を述べたが、緊張度、興奮度、冷静度等の様々なものと考えられる。

【0058】また、抽出する感性の種類は、平常と非常に2種類でも良い。抽出する感性の種類は、音声制御装置と組み合わされる出力装置5の操作性・安全性・耐久性・用途・設置場所等を考慮して、適宜設定されるものである。また、音声特徴のレベルを3段階としたが、本願はこれに限定されるものではない。

【0059】また、本実施例の出力装置5としては、テレビジョン、照明装置、情報処理装置の事例で説明した

が、本発明を実施する場合には、オン・オフ制御または可変制御する機器または応答を制御する機器であればよく、その用途が特定されるものではない。また、音声制御装置と出力装置とは、通常一体化されて販売される。

【0060】また、感性レベル分析部36での感性レベルの分析及び出力制御手段4での制御情報の決定には、ファジイ理論を用いるようにしてもよい。また、上記実施例では、音声特徴レベル分析部36には、発生速度検出部31、音声ピッチ周波数検出部32、音量検出部33の出力を入力しているが、本発明を実施する場合には、更に、周波数成分を計測した音声スペクトル情報（音韻「音色」）の検出を行い、その出力も使用してもよい。

【0061】図4を参照しつつ、本発明の第2実施例を説明する。なお、図中、図1と同一部分には、同一符号を付して重複説明を省略する。上記第1実施例では、音声制御装置の特に感情による制御について説明したが、この第2実施例は、話者の感情だけでなく、話者の状態を検出している。ここでいう状態とは、「喉頭癌や風邪などの体調状態」や「鼻をつまんで喋っている」や「マイクに口をつけて喋っている」などである。

【0062】つまり、上記の話者の状態を検出するように、感性レベル分析用知識ベース格納部に感情を検出するためだけなく、体調等を検出するためのデータを予め追加格納して状態レベル分析用知識ベース格納部37'とする。同様に、出力制御用知識ベース格納部42'には、感情に対応するだけなく、体調等にも対応するデータを予め追加格納する。

【0063】図4において、音声スペクトル検出部39は、ある時間内の音声データをDFT（デジタルフーリエ変換）分析して、音声スペクトルを得ている。普通、穏やかな話し方では、音声波形は正弦波に近く、あまり高調波成分はない。しかし、怒鳴ったり強く言うときは、三角波に近くなり、高調波成分が多く含まれる。

【0064】基準音声特徴量格納部35'には、音声分析する場合の基準となる音声特徴量（発生速度、音声ピッチ周波数、音量、音声スペクトル）が記憶されている。音声特徴レベル判定基準格納部34'には、レベル分類の【発生速度レベル基準】と【音声ピッチ周波数レベル基準】、【音量特徴レベル基準】、【音声スペクト

ルレベル基準】が格納されている。

【0065】音声特徴レベル分析部36'は、検出部31, 32, 33, 39で得られた種々の音声特徴量と基準音声特徴量とを比較し、この比較結果を音声特徴レベル判定基準格納部34'に格納されているレベル基準と比較する。比較した結果を状態レベル分析38'に送出する。状態レベル分析38'ではその結果を基に状態レベル分析用知識ベース格納部37'に格納されている知識ベースに基づいて話者の状態レベル（状態と程度）を決定する。

【0066】出力制御手段4は、状態レベルと音声コードを基に、状態レベル分析用知識ベース格納部37'に格納されている知識ベース（条件マップ）に基づいて、出力装置5への制御情報を決定する。例えば、対話型のシステムでは、話者の体調が異常であれば、状態レベル分析部38'でその状態を判断し、その結果を基に出力制御手段4で応答内容や応答メッセージを変えることも考えられる。

【0067】例としては、話者の声「始動」が、風邪気味の声であると状態情報抽出手段3'で検出すると、出力制御手段4は、出力装置5を始動させる制御情報を出力すると共に、「病気ですか？」のメッセージを画面出力する。そして、次に、話者の声「違う」又は、他の命令等（話者がこのメッセージを無視）が入力されると、出力制御手段4は、通常通りに動作する。

【0068】このときに、話者の声「ハイ」「そうだ」「うるさい」等が、入力されると、出力制御手段4は、出力制御用知識データベース格納部42'で使用する条件マップを体調不調話者用の条件マップに変更する。この体調不調話者用の条件マップに変更した後に、話者の声「業務ファイルを開け」が入力されると、出力制御手段4は、業務ファイルを開く制御情報を出力すると共に、「大丈夫ですか？」のメッセージを画面出力する。

【0069】また、この体調不調話者用の条件マップに変更した後に、話者の声「ゲームファイルを開け」が入力されると、出力制御手段4は、ゲームファイルを開く制御情報を出力せずに、「具合が悪いのだから遊んだりしちゃダメ！」のメッセージを画面出力させる。尚、上記実施例では、状態情報抽出手段3と音声コード認識手段2は、別ブロックで表したが、これは、当然、一体化してもよい。

【0070】又、上記実施例では、状態情報抽出手段3'は、音声が入力される度に、常に状態レベル情報を出力するが、これは、状態レベルが変化した時だけ出力するようにしても良い。又、抽出精度が低い時には、出

力しないようにしてもよい。また、状態レベル情報の精度が悪い時には、前回検出した状態レベル情報を出力してもよい。また、状態レベル情報の精度が悪い時には、平常の状態レベル情報を出力してもよい。また、状態レベル情報の精度が悪い時には、非常の状態レベル情報を出力してもよい。また、状態情報抽出手段3'は、常に、音声コードと音声特徴レベルの両方により、状態レベル情報を決定したが、全ての場合に、両者が必要というわけでもない。同様に、出力制御手段4', 3'は、常に、音声コードと状態レベルの両方により、制御情報を決定したが、全ての場合において、音声コードと状態レベルの両情報が常に必須というわけでもない。この両情報の必要性は、音声制御装置と組み合わされる出力装置5の操作性・安全性・耐久性・用途・設置場所等を考慮して、適宜設定されるものである。

【0071】

【発明の効果】本発明によれば、音声制御装置は、話者からの音声命令に対して、額面通りの命令内容でなく、その音声命令に含まれるニュアンスにも対応した制御が可能となる。更に、本発明によれば、音声制御装置は、話者からの音声命令の入力時に状態を抽出出来なくても適切な制御を行うことが出来る。

【図面の簡単な説明】

【図1】本発明の第1実施例の音声制御装置を示すブロック図である。

【図2】この第1実施例の音声制御装置の処理動作を説明する説明図である。

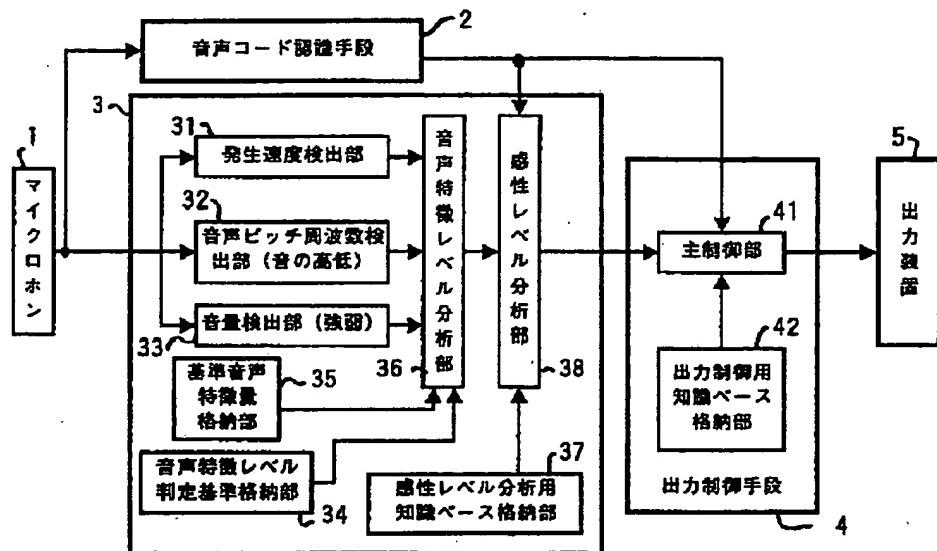
【図3】第1実施例の音声制御装置の状態情報抽出手段が行う制御処理を示すフローチャートである。

【図4】本発明の第2実施例の音声制御装置を示すブロック図である。

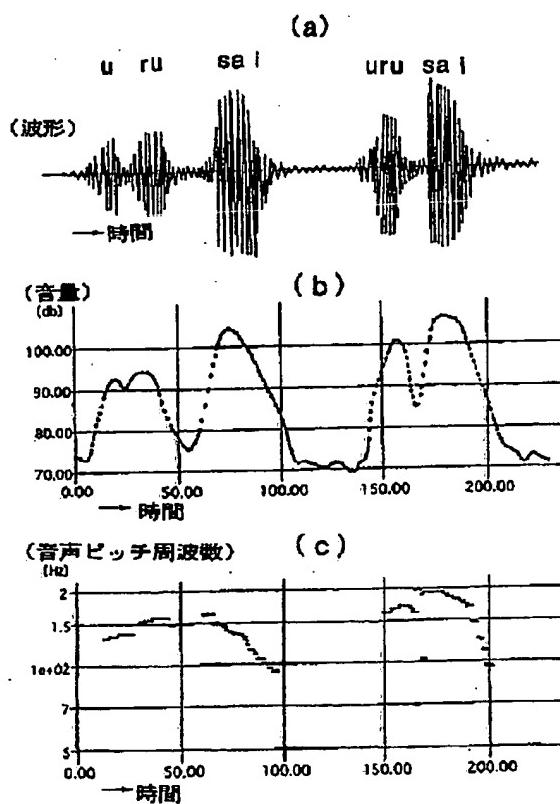
【符号の説明】

- 1・・・マイクロホン、
- 2・・・音声コード認識手段（コード認識手段）、
- 3・・・状態情報抽出手段、
- 3'・・・状態情報抽出手段、
- 4・・・出力制御手段、
- 36・・・音声特徴レベル分析部、
- 36'・・・音声特徴レベル分析部、
- 37・・・感性レベル分析用知識ベース格納部、
- 37'・・・状態レベル分析用知識ベース格納部、
- 38・・・感性レベル分析部、
- 38'・・・状態レベル分析部、
- 41・・・主制御部、
- 42・・・出力制御用知識ベース格納部、
- 42'・・・出力制御用知識ベース格納部。

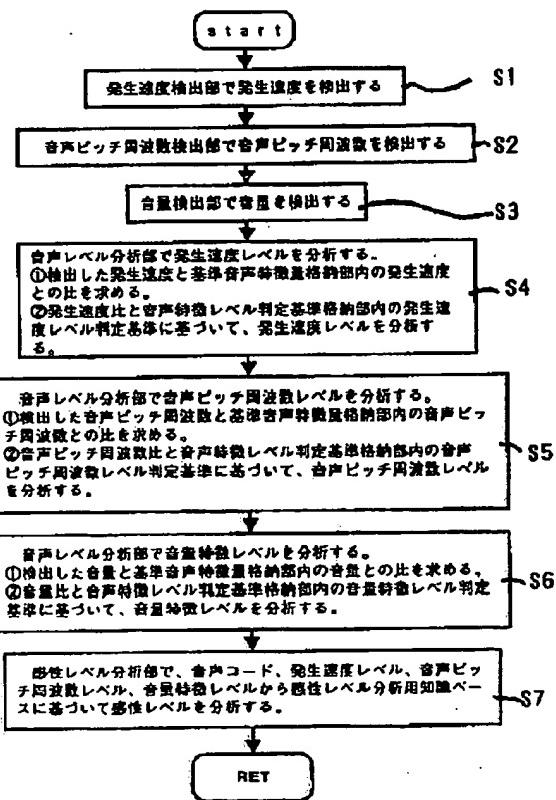
【図1】



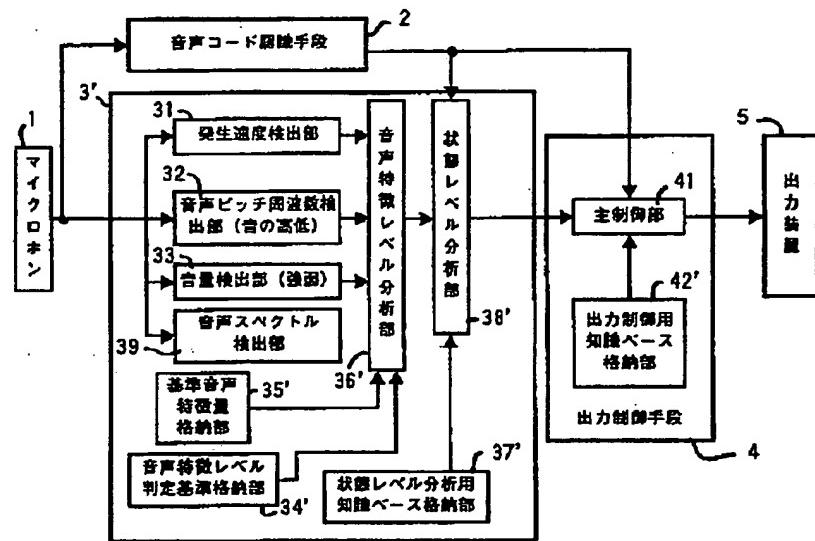
【図2】



【図3】



【図 4】



PATENT ABSTRACTS OF JAPAN

(11)Publication number : **10-055194**
(43)Date of publication of application : **24.02.1998**

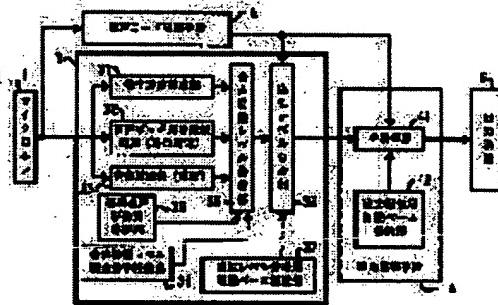
(51)Int.CI. G10L 3/00
G10L 3/00
G10L 3/00
G10L 3/00

(54) DEVICE AND METHOD OF VOICE CONTROL

(57)Abstract:

PROBLEM TO BE SOLVED: To realize control in which human feeling, etc., are taken into account and nuance included in a vocal command is also corresponded by controlling the output device based on the outputs of the means which recognizes the code information corresponding to inputted vocal information and the means which extracts the condition information of the accompanying speaker.

SOLUTION: A voice code recognizing means 2 recognizes the code information corresponding to the inputted vocal information, and a condition information extracting means 3 extracts the condition information (feeling, physical conditions, etc.) of the speaker accompanying this vocal information. And an output control means 4 controls an output device 5 based on the output of the means 2 and that of the means 3. In this case, it is desirable that, when the condition information of the speaker cannot be detected, control is performed as the condition detected immediately before or as the normal or emergent condition. In compliance with the code information corresponding to vocal information and the condition of the speaker it is desirable to output a message which cares for this speaker from the output device 5.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than
the examiner's decision of rejection or
application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision
of rejection]

[Date of requesting appeal against examiner's
decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

BLACK BORDERS

IMAGE CUT OFF AT TOP, BOTTOM OR SIDES

FADED TEXT OR DRAWING

BLURRED OR ILLEGIBLE TEXT OR DRAWING

SKEWED/SLANTED IMAGES

COLOR OR BLACK AND WHITE PHOTOGRAPHS

GRAY SCALE DOCUMENTS

LINES OR MARKS ON ORIGINAL DOCUMENT

REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.